SUPPLEMENTARY MATERIALS

Supplementary Methods

Clinical Cohort Study Population

All 8,753 TNBC cases from a cohort of 140,449 individuals subjected to clinical germline cancer panel testing between March 2012 and June 2016 at a clinical testing laboratory (Ambry Genetics-Aliso Viejo, CA) were included in this study. Demographic, clinical history, and family history of cancer information was collected from test requisition forms, clinic notes, and pedigrees provided by ordering clinicians at the time of testing. Information was collected on current age, personal history and age at diagnosis of all cancers, ancestry, family history of cancer with cancer type, and age at diagnosis among relatives. Family history was limited to first and second-degree relatives. Families with breast or ovarian cancer in two or more individuals, on the same parental side, were considered positive for family history for each cancer. TNBC status, defined as ER-negative, PR–negative and HER2–negative breast cancer was provided by ordering clinicians and/or clinical tumor pathology reports. ER, PR, and HER2 status was available on 50,820 of 74,649 (67.3%) breast cancer cases.

To assess data quality, a review of a random sample of 1200 (10%) breast and ovarian cancer patient intake forms was conducted. Of these, 43.3% (520 of 1200) with additional clinical history documentation available (clinic notes, pedigrees, detailed letters of medical necessity). Consistent information was observed for 100% (429/429) of breast cancers, 99.5% (409/411) for age at breast cancer diagnosis, 99.7% (331/332) for breast tumor pathology, 97.9% (275/281) for breast tumor hormone receptors status, 100% (90/90) for ovarian cancer, and 100% (88/88) for age at ovarian cancer diagnosis, suggesting that the intake data for the cohort is of reasonably high quality.

Triple-Negative Breast Cancer Consortium (TNBCC) study

The TNBCC study included 2,148 TNBC cases from 12 clinical centers in the USA (MCBCS, DFCI, OSU, RPCI, KUMC and FCCC), Germany (BBCC and GENICA), Finland (HEBCS), Greece (DEMOKRITOS) and the United Kingdom (UK) (SBCS); and the POSH UK study of women diagnosed under age 40. Selection of TNBC cases was independent of family history of breast and ovarian cancer. All 2,148 TNBC cases were recruited to institutional review board approved studies. Information on age of diagnosis, family history of cancer, and self-reported ethnicity was provided by participating studies. Family history was recorded as all first and second-degree relatives with breast or ovarian cancer. TNBC cases in TNBCC were individuals with ER-negative, PR–negative and HER2–negative (0-1 by immunohistochemistry) breast cancer. Because a subset of patients were recruited from 1995-2011, ER and PR-negative status was defined as <10% of the tumor cells stained.

Multigene Panel Testing for the Clinical Cohort

Mutation testing was performed by sequencing of targeted custom capture products from several clinical multigene panels and targeted chromosomal microarray analysis. Genomic deoxyribonucleic acid (gDNA) was isolated from the patient's blood or saliva specimen using a standardized methodology (Qiagen, Valencia, CA). Sequence enrichment was performed by incorporating the gDNA onto microfluidics chip or into microdroplets along with primer pairs or by a bait-capture methodology using long biotinylated oligonucleotide probes (RainDance Technologies, Billerica, MA or Integrated DNA Technologies, San Diego, CA), followed by PCR

and then NGS analysis (Illumina, San Diego, CA) of all coding exons plus at least five bases into the 5' and 3' ends of all the introns and untranslated regions (5'UTR and 3'UTR). A targeted chromosomal microarray was used for the detection of gross deletions and duplications for all genes except *PMS2* (Agilent, Santa Clara, CA). Gross deletion/duplication analysis of *PMS2* was performed using MLPA kit# P008-B1 (MRC-Holland, Amsterdam, Netherlands) and Sanger sequencing. Initial data processing and base calling was done using RTA 1.12.4 (HiSeq Control Software 1.4.5; Illumina). Sequence quality filtering at Q20 was executed with the CASAVA software (version 1.8.2; Illumina, Hayward, CA). Sequence fragments were aligned to the reference human genome (GRCh37), and variant calls were generated using CASAVA. Variants were annotated with the Ambry Variant Analyzer, a proprietary alignment and variant annotation software (Ambry Genetics). All variants identified by Ambry Genetics are submitted to the ClinVar public database.

Panel-based Mutation Analysis for TNBCC

Germline DNA samples from 2,148 TNBC cases were subjected to custom capture (Agilent eArray) of all coding sequences and intron/exon boundaries of coding exons of 17 breast cancer predisposition genes (*BRCA1, BRCA2, PALB2, BARD1, BRIP1, RAD51C, RAD51D, RAD50, NBN, MRE11A, XRCC2, ATM, CHEK2, TP53, PTEN, STK11, CDH1*) as part of a panel of 122 DNA repair genes. Products from each capture reaction were sequenced on a HiSeq 2000. Paired end reads (100bp) were aligned to the hg19 reference human genome using Novoalign (Novocraft Technologies, Malaysia). Realignment and recalibration was performed using GATK (VN:1.6-7). Germline variations were called with a combination of GATK UnifiedGenotyper and Samtools(VN 0.1.18). All likely deleterious mutations were validated by Sanger sequencing.

Statistical Methods

The frequency of all pathogenic variants (PVs) within each gene in TNBC cases of Caucasian ancestry was compared with the frequency of pathogenic mutations in the ExAC_NFE_non-TCGA (ExAC non-Finnish European, excluding The Cancer Genome Atlas Project (TCGA) cases) reference controls after data cleaning and filtering. Copy Number Variants in all genes and variants in pseudogene homology regions (*PMS2* Exon 9, 11-15) were excluded from cases and controls for risk estimation because these alterations were not individually validated in ExAC or gnomAD. Established low penetrance mutations (*APC* p.Ile1307Lys; *CHEK2* p.Ile157Thr) were excluded. Suspected mosaic somatic variants (allele ratio>70:30), variants with minor allele frequency (MAF)>0.3%, and low penetrance variants were excluded from both cases and controls (e.g., CHEK2 p.Ile157Thr). Truncating variants in the last 55bp of the penultimate exon or the last exon that potentially avoided nonsense mediated mRNA decay, and did not disrupt a known functional domain, were excluded. Large genomic rearrangements of one or more exons were excluded from case-control comparisons because rearrangements were not validated among reference controls.

Associations between combined PVs in each gene and TNBC for the Caucasian population were estimated by odds ratio (OR) and corresponding 95% confidence intervals based on Fisher's exact test. P-values <0.05 were considered statistically significant. Sensitivity analyses for associations were performed for all races and ethnicities combined; by restricting to ExAC PASS reference controls; TNBCs as the first cancer, and when excluding personal or family history of ovarian or colorectal cancer. Similar studies for Caucasian TNBCs were conducted using a combined gnomAD_NFE (non-Finnish European) and gnomAD Ashkenazi Jewish (ASJ) PASS reference control dataset. While the gnomAD controls partially overlap with ExAC_NFE_non-TCGA controls, the substantially increased number along with updated variant calling algorithms identified gnomAD as an independent reference control dataset. The frequency of PVs by gene in African American TNBC patients from the clinical cohort and

African and African American reference controls from gnomAD_AFR were also compared. gnomAD controls were used instead of ExAC controls because of limited numbers of PVs in the ExAC_AFR control dataset. All tests were two-sided. Associations between PVs and age of PC diagnosis were evaluated using the Kolmogorov-Smirnov test. Enrichment of mutations by gene in TNBC relative to all non-TNBC breast cancers was estimated by logistic regression, with adjustment for age of diagnosis and family history of breast cancer.

Absolute Risk Estimation for Breast Cancer and TNBC from the Clinical Cohort

Absolute risk is the probability an individual with a measured set of risk factors (e.g. mutation status, a Polygenic Risk Score (PRS), and family history of disease) and is disease free at age a will be diagnosed with the disease in the subsequent τ years[1]. Let Z be the set of measured risk factors, then we can express the absolute risk as:

$$R(a,\tau,Z) = \int_a^{a+\tau} h_1(u|Z) \times \exp\left(-\int_a^u \{h_1(v|Z) + h_2(v|Z)\}\,\mathrm{d}v\right)\mathrm{d}u$$

where $h_1(a|Z)$ is the conditional disease-specific hazard at age a and $h_2(a|Z)$ is the competing risks hazard at age a. The competing risks may include other diseases or death. The hazards can be parameterized as:

$$h_1(a|Z) = h_{10}(a)\exp(\beta Z)$$

and

$$h_2(a|Z) = h_{20}(a)\exp(\gamma Z)$$

where $h_{i0}(a)$ is the baseline hazard and β and γ represent the relative risks for each risk factor in Z for the disease of interest and competing risks, respectively. Estimates for β and γ can be estimated from the case-control data. The baseline hazards can be estimated by the following relationship between the baseline hazard and the marginal hazard

$$h_1^*(a) = h_{10}(a)\mathrm{E}(\exp(\beta Z)) \approx \int h_{10}(a)\exp(\beta z)dF(z)$$

where $F(Z)$ denotes the distribution of the risk factors in the population. For each age, we solve for $h_{10}(a)$ by using the expected distribution of the risk factor in the population[2,3].

We estimated the triple negative breast cancer (TNBC) absolute risk for a women at age *a* without a diagnosis with any breast cancer, and the risk factor of interest being gene specific mutation carrier status, the odds ratio estimates from the Ambry cases versus ExAC controls were used as estimates for β, and SEER TNBC specific incidence rates[4] were utilized for the estimation of the baseline hazard combined with the ExAC mutation frequency for the population frequency. The age-specific competing hazards model, $h_2(a|Z)$, involves a mixture of the hazard for HR positive breast cancer, HER2-enriched breast cancer, and non-breast cancer death. The three hazard functions were estimated (data not shown), with the first two utilizing subtype specific odds ratios from the clinical cohort combined with subtype specific SEER incidence rates, and the third utilizing US mortality data.

Inclusion criteria for absolute risk estimation

Female breast cancer patients with available tumor estrogen receptor (ER), progesterone receptor (PR), and HER2 histopathology data were used for estimation of breast cancer subtype specific risks for pathogenic variants in each gene. Out of 73,570 female breast cancer patients,

a total of 8,553 TNBC, 38,171 ER-positive, and 2,332 ER-negative/HER2-positive cases were included in analysis. For subtype specific risk estimation, patients previously tested for BRCA1/2 mutations, or with multiple pathogenic mutations, copy number variation (CNV), or PMS2 pathogenic variants in exons common with pseudogenes were excluded.

References

1. Gail MH, Brinton LA, Byar DP, et al. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *J Natl Cancer Inst.* 1989;81(24):1879-1886.

2. Costantino JP, Gail MH, Pee D, et al. Validation studies for models projecting the risk of invasive and total breast cancer incidence. *J Natl Cancer Inst.* 1999;91(18):1541-1548.

3. Chatterjee N, Shi J, Garcia-Closas M. Developing and evaluating polygenic risk prediction models for stratified disease prevention. *Nat Rev Genet.* 2016;17(7):392-406.

4. National Cancer Institute, DCCPS, Surveillance Research Program, Surveillance Systems Branch. Surveillance, Epidemiology, and End Results (SEER) Program Research Data (1973-2014). https://www.seer.cancer.gov. April 2017 (Based on November 2016 Submission).

Supplementary Tables

Supplementary Table 1: Population characteristics of clinically tested breast cancer cases

Characteristics	Clinical TNBC					
	Caucasian			African American		
	n	%	% PV	n	%	% PV
Total patients	5498	100	14.0	1271	100	14.6
Gender						
Female	5492	99.9	14.0	1270	99.9	14.6
Male	5	0.1	ND	1	0.1	ND
Unknown	1	0.0	ND	0	0.0	ND
Race/Ethnicity						
African-American	0	0.0	NA	1271	100.0	14.6
Ashkenazi-Jewish	292	5.3	14.3	0	0.0	NA
Asian	0	0.0	NA	0	0.0	NA
Caucasian	5206	94.7	14.0	0	0.0	NA
Hispanic	0	0.0	NA	0	0.0	NA
Other/Unknown	0	0.0	NA	0	0.0	NA
Personal History of Cancer						
Breast						
Age at diagnosis, yr						
≤ 36	671	12.2	18.9	180	14.2	26.3
≤ 45	1772	32.2	17.0	487	38.3	20.7
≤ 50	2648	48.2	16.1	709	55.8	18.8
≤ 60	4454	81.0	14.7	1105	86.9	15.0
> 60	1022	18.6	11.4	162	12.7	11.4
Unknown	22	0.4	ND	4	0.3	ND
Multiple breast cancer	881	16.0	17.6	196	15.4	24.5
Ovarian	79	1.4	ND	9	0.7	ND
Colorectal	63	1.1	ND	11	0.9	ND
Pancreatic	21	0.4	ND	1	0.1	ND
Family History of Cancer*		0.0				
Breast (no ovarian)	2584	47.0	14.1	610	48.0	17.1
Breast & Ovarian	396	7.2	25.7	63	5.0	ND
Ovarian (no breast)	273	5.0	23.5	47	3.7	ND
Colorectal	1340	24.4	13.1	200	15.7	15.3
Pancreatic	517	9.4	16.0	95	7.5	ND
No Breast, Ovarian, Colorectal or pancreatic	1292	23.5	10.6	338	26.6	10.1
Mean age at TNBC diagnosis	Mean ± SD	Range		Mean ± SD	Range	
	50 ±11.4	18-90		48.6 ±10.8	19-83	

ND = not determined, SD = Standard deviation; PV = pathogenic variant
*1st and 2nd degree relatives with the relevant cancer

Supplementary Table 2. TNBC gene-based variant frequency by race and ethnicity in the clinical and TNBCC studies

	Clinical Cohort									TNBCC Cohort					
	All race/ethnicity			Caucasian			African American			All race/ethnicity			Caucasian		
	Mutated alleles	Cases	Freq (%)	Mutated alleles	Cases	Freq (%)	Mutated alleles	Cases	Freq (%)	Mutated alleles	Cases	Freq (%)	Mutated alleles	Cases	Freq (%)
TNBC predisposition genes															
BARD1	48	6464	0.74	27	4144	0.65	11	886	1.24	10	2148	0.47	10	2095	0.48
BRCA1	513	8537	6.01	297	5332	5.57	74	1257	5.89	166	2148	7.73	164	2095	7.83
BRCA2	201	8537	2.35	119	5332	2.23	39	1257	3.10	58	2148	2.70	54	2095	2.58
BRIP1	27	6464	0.42	18	4144	0.43	4	886	0.45	10	2148	0.47	10	2095	0.48
PALB2	111	6980	1.59	75	4441	1.69	8	964	0.83	22	2148	1.02	22	2095	1.05
RAD51C	31	6464	0.48	20	4144	0.48	5	886	0.56	8	2148	0.37	8	2095	0.38
RAD51D	16	6095	0.26	10	3868	0.26	3	859	0.35	8	2148	0.37	8	2095	0.38
TP53	14	8741	0.16	10	5490	0.18	3	1270	0.24	2	2148	0.09	2	2095	0.10
Total frequency (%)			12.01			11.49			12.66			13.22			13.28
Other cancer predisposition genes															
ATM	17	6652	0.26	13	4265	0.30	1	913	0.11	4	2148	0.19	4	2095	0.19
CDH1	5	8505	0.06	1	5329	0.02	3	1233	0.24	ND	ND	ND	ND	ND	ND
CDKN2A	5	1790	0.28	3	1213	0.25	-	192	ND	ND	ND	ND	ND	ND	ND
*CHEK2**	22	6639	0.33	19	4252	0.45	-	913	ND	2	2148	0.09	2	2095	0.10
MLH1	5	3497	0.14	4	2301	0.17	-	407	ND	ND	ND	ND	ND	ND	ND
MRE11A	6	6464	0.09	2	4144	0.05	3	886	0.34	4	2148	0.19	4	2095	0.19
MSH2	3	3497	0.09	2	2301	0.09	-	407	ND	0	372		0	372	
MSH6	9	3497	0.26	7	2301	0.30	-	407	ND	1	372	0.27	1	372	0.27
NF1	9	6097	0.15	9	3870	0.23	-	859	ND	ND	ND	ND	ND	ND	ND
NBN	12	6464	0.19	10	4144	0.24	1	886	0.11	2	2148	0.09	2	2095	0.10
PMS2	9	3497	0.26	4	2301	0.17	3	407	0.74	ND	ND	ND	ND	ND	ND
PTEN	4	8719	0.05	2	5471	0.04	1	1269	0.08	1	2148	0.05	1	2095	0.05
RAD50	14	6464	0.22	9	4144	0.22	3	886	0.34	6	2148	0.28	5	2095	0.24
XRCC2	ND	ND	ND	ND	ND	ND	ND	ND	ND	2	2148	0.09	2	2095	0.10
Total frequency (%)			2.38			2.53			1.96			1.25			1.24

(-) no mutated alleles; Freq = variant frequency in percent; ND = not determined

* Excluding *CHEK2* p.Ile157Thr and p.Ser428Phe

Supplementary Table 3. Gene-based variant frequency by personal and family history of cancer in the TNBC clinical cohort

	Overall TNBC cases			Personal history of bilateral cancer			1st or 2nd degree family history of cancer					
							Breast			Ovarian		
	Mutation carriers	No. cases	Freq. (%)	Mutation carriers	No. cases	Freq. (%)	Mutation carriers	No. cases	Freq. (%)	Mutation carriers	No. cases	Freq. (%)
Breast cancer predisposition genes												
ATM	17	6652	0.26	5	1037	0.48	8	3091	0.26	3	295	1.02
BARD1†	48	6464	0.74	9	1010	0.89	27	3016	0.90	1	288	0.35
BRCA1†	513	8537	6.01	113	1264	8.94	250	3858	6.48	35	394	8.88
BRCA2†	201	8537	2.35	41	1264	3.24	108	3858	2.80	20	394	5.08
CHEK2	22	6639	0.33	4	1034	0.39	9	3087	0.29	1	295	0.34
PALB2†	111	6980	1.59	20	1076	1.86	64	3249	1.97	4	306	1.31
PTEN	4	8719	0.05	1	1305	0.08	3	3971	0.08	-	402	ND
RAD51D†	16	6095	0.26	2	928	0.22	6	2804	0.21	2	275	0.73
TP53†	14	8741	0.16	6	1307	0.46	6	3978	0.15	-	402	ND
Total frequency (%)			11.75			16.55			13.14			17.70
Other cancer predisposition genes												
*BRIP1**	27	6464	0.42	5	1010	0.50	10	3016	0.33	2	288	0.69
CDH1	5	8505	0.06	2	1267	0.16	4	3891	0.10	-	379	ND
CDKN2A	5	1790	0.28	2	280	0.71	1	827	0.12	3	82	3.66
MLH1	5	3497	0.14	2	544	0.37	3	1492	0.20	1	237	0.42
MRE11A	6	6464	0.09	2	1010	0.20	2	3016	0.07	-	288	ND
MSH2	3	3497	0.09	-	544	ND	1	1492	0.07	-	237	ND
MSH6	9	3497	0.26	2	544	0.37	4	1492	0.27	-	237	ND
NBN	12	6464	0.19	1	1010	0.10	7	3016	0.23	1	288	0.35
NF1	9	6097	0.15	2	928	0.22	4	2805	0.14	-	275	ND
PMS2	9	3497	0.26	-	544	ND	1	1492	0.07	3	237	1.27
RAD50	14	6464	0.22	1	1010	0.10	6	3016	0.20	3	288	1.04
*RAD51C**	31	6464	0.48	3	1010	0.30	13	3016	0.43	2	288	0.69
Total frequency (%)			2.62			3.01			2.23			8.12

(-) no mutated alleles; Freq = variant frequency in percent; ND = not determined

* Established TNBC predisposition genes

Supplementary Table 4. Estimated risks of TNBC in Caucasian patients associated with mutations in non-TNBC hereditary cancer panel genes

Non-TNBC predisposition gene	Clinical Cohort		TNBCC Cohort		ExAC controls		Clinical TNBC Risk				TNBCC TNBC Risk			
	Mutated Alleles	No. Cases	Mutated Alleles	No. Cases	Mutated Alleles	No. Controls	OR	95%CI lower	95%CI upper	p-value*	OR	95%CI lower	95%CI upper	p-value*
ATM	13	4210	4	2003	102	26644	0.81	0.43	1.45	0.59	0.52	0.18	1.36	0.25
CDH1†	1	5263	ND	ND	3	25961	1.64	0.06	14.91	0.52	ND	ND	ND	ND
CDKN2A†	3	1195	ND	ND	9	24312	6.79	1.58	24.22	0.02	ND	ND	ND	ND
CHEK2	17	4197	2	2003	232	25215	0.44	0.26	0.73	3.63×10^{-4}	0.11	0.02	0.40	6.75×10^{-6}
MLH1	3	2266	ND	ND	10	26639	3.53	0.83	12.72	0.08	ND	ND	ND	ND
MRE11A	2	4090	4	2003	25	26767	0.52	0.09	2.03	0.57	2.14	0.68	5.96	0.14
MSH2†	1	2266	0	372	6	25329	1.86	0.08	13.90	0.45	ND	ND	ND	ND
MSH6	7	2266	1	372	34	26151	2.38	1.03	5.33	0.04	2.07	0.10	12.67	0.39
NBN	10	4090	2	2003	41	26265	1.57	0.75	3.19	0.22	0.64	0.11	2.48	0.77
NF1	9	3816	NA	NA	29	26131	2.13	0.99	4.46	0.05	NA	NA	NA	NA
PMS2†	3	2266	ND	ND	32	26230	1.09	0.28	3.38	0.76	ND	ND	ND	ND
PTEN†	2	5404	1	2003	1	24166	8.95	0.70	259.36	0.09	12.07	0.31	464.53	0.15
RAD50	7	4090	4	2003	58	26474	0.78	0.35	1.68	0.72	0.91	0.30	2.50	1.00
XRCC2†	ND	ND	2	2003	14	27085	ND	ND	ND	ND	1.93	0.31	8.01	0.30

ND: not determined; OR = Odds ratio; CI = confidence interval; TNBCC = Triple Negative Breast Cancer Consortium

* Statistical significance of associations were estimated using Fisher's exact test. All tests were two-sided.

†Insufficient events for estimation of risk

Supplementary Table 5. Estimated risks of TNBC in Caucasian patients using gnomAD controls

Gene	Clinical Cohort		TNBCC Cohort		gnomAD controls		Clinical TNBC Risk				TNBCC TNBC Risk			
	Mutated Alleles	No. Cases	Mutated Alleles	No. Cases	Mutated Alleles	No. Controls	OR	95%CI lower	95%CI upper	p-value*	OR	95%CI lower	95%CI upper	p-value*
TNBC associated genes														
BARD1	25	4090	9	2003	56	59601	6.52	3.99	10.46	$2.27x10^{-11}$	4.79	2.31	9.67	$2.49 x10^{-4}$
BRCA1	255	5265	158	2003	194	60628	15.49	12.80	19.11	$1.05x10^{-16}$	25.62	20.67	31.79	$<2.2x10^{-16}$
BRCA2	115	5265	51	2003	231	60033	5.73	4.56	7.21	$8.52x10^{-41}$	6.69	4.90	9.09	$<2.2x10^{-16}$
BRIP1	17	4090	9	2003	119	60616	2.12	1.23	3.55	$6.98x10^{-3}$	2.29	1.12	4.46	0.02
PALB2	70	4383	17	2003	98	60675	9.96	7.23	13.65	$2.13x10^{-37}$	5.27	3.07	8.86	$1.63 x10^{-7}$
RAD51C	15	4090	8	2003	51	60657	4.37	2.42	7.74	$1.23x10^{-5}$	4.76	2.06	9.83	$5.62x10^{-4}$
RAD51D	8	3814	7	2003	28	60537	4.54	1.89	10.51	$1.03x10^{-3}$	7.57	3.20	17.78	$1.05 x10^{-4}$
TP53	10	5423	2	2003	20	60671	5.60	2.41	11.99	$8.96x10^{-5}$	3.03	0.51	12.44	0.16
TP53†	6	1055	2	504	20	60671	17.30	6.80	42.26	$4.25x10^{-6}$	12.06	2.01	49.65	0.01
Other predisposition genes														
ATM	13	4210	4	2003	236	60572	0.79	0.43	1.39	0.52	0.51	0.17	1.36	0.27
CDH1‡	1	5263	ND	ND	8	59483	1.41	0.06	9.01	0.53	ND	ND	ND	ND
CDKN2A‡	3	1195	ND	ND	13	56936	11.01	2.68	36.95	$3.98 x10^{-3}$	ND	ND	ND	ND
CHEK2	17	4197	2	2003	488	59768	0.50	0.29	0.80	$2.12 x10^{-3}$	0.12	0.02	0.45	$2.96x10^{-5}$
MLH1	3	2266	ND	ND	12	60372	6.66	1.61	23.26	0.02	ND	ND	ND	ND
MRE11A	2	4090	4	2003	54	60581	0.55	0.09	2.07	0.58	2.24	0.74	6.19	0.12
MSH2‡	1	2266	0	372	14	60246	1.90	0.09	11.50	0.43	ND	ND	ND	ND
MSH6	7	2266	1	372	67	59898	2.76	1.25	6.20	0.02	2.41	0.12	13.79	0.34
NF1	9	3816	ND	ND	23	60462	6.21	2.85	13.60	$7.31 x10^{-5}$	ND	ND	ND	ND
NBN	10	4090	2	2003	96	60497	1.54	0.78	2.91	0.23	0.63	0.11	2.28	0.77
PMS2‡	3	2266	ND	ND	58	59272	1.35	0.35	4.15	0.49	ND	ND	ND	ND
PTEN‡	2	5404	1	2003	4	60242	5.57	0.75	29.94	0.08	7.52	0.31	57.74	0.15
RAD50	7	4090	4	2003	143	60466	0.72	0.33	1.51	0.50	0.84	0.29	2.17	1.00

ND = not determined or no PVs; OR = Odds ratio; CI = confidence interval; TNBCC = Triple Negative Breast Cancer Consortium

* Statistical significance of associations were estimated using Fisher's exact test. All tests were two-sided.

†Age at diagnosis of 40 years or younger

‡ Insufficient events for estimation of risk

Supplementary Table 6: Associations with TNBC by gene using ExAC NFE-nonTCGA PASS reference controls

Gene	Clinical cohort		ExAC controls		Cancer Risk			
	Mutated Alleles	No. Cases	Mutated Alleles	No. Controls	OR	95%CI lower	95%CI upper	p-value*
ATM	13	4210	93	26715	0.89	0.47	1.60	0.78
BARD1	25	4090	20	26193	8.03	4.47	14.81	3.53×10^{-11}
BRCA1	255	5265	69	26913	19.33	14.82	25.41	$<2.2 \times 10^{-16}$
BRCA2	115	5265	94	26804	6.29	4.75	8.31	$<2.2 \times 10^{-16}$
BRIP1	17	4090	41	26864	2.73	1.52	4.81	1.28×10^{-3}
CDH1	1	5263	2	26409	2.51	0.09	32.06	0.42
CDKN2A	3	1195	8	24424	7.67	1.75	0.17	0.01
CHEK2	17	4197	227	25296	0.45	0.27	0.74	6.23×10^{-4}
MLH1	3	2266	10	26644	3.53	0.83	12.72	0.08
MRE11A	2	4090	22	26784	0.60	0.10	2.38	0.76
MSH2	1	2266	6	25463	1.87	0.08	13.97	0.45
MSH6	7	2266	28	26419	2.92	1.23	6.85	0.02
NBN	10	4090	39	26282	1.65	0.78	3.27	0.15
NF1	9	3816	17	26508	3.68	1.61	8.23	3.28×10^{-3}
PALB2	70	4383	26	26871	16.63	10.53	26.40	$<2.2 \times 10^{-16}$
PMS2	3	2266	24	26304	1.45	0.37	4.81	0.47
PTEN	2	5404	1	24619	9.11	0.71	264.22	0.09
RAD50	7	4090	52	26556	0.87	0.39	1.91	0.85
RAD51C	15	4090	34	26697	2.88	1.55	5.32	1.27×10^{-3}
RAD51D	8	3814	8	26576	6.97	2.61	18.67	3.08×10^{-4}
TP53	10	5423	16	26757	3.09	1.33	6.76	7.10×10^{-3}
TP53†	6	1024	16	26757	9.82	3.77	26.26	1.12×10^{-4}

OR = Odds ratio; CI = confidence interval

* Significance of associations were estimated using Fisher's exact test. All tests were two-sided.

†Age at diagnosis of 40 years or younger

Supplementary Table 7: Associations with TNBC among Caucasian with TNBC as first cancer diagnosed

Gene	Clinical cohort		ExAC controls		Cancer Risk			
	Mutated Alleles	No. Cases	Mutated Alleles	No. Controls	OR	95%CI lower	95%CI upper	p-value*
ATM	10	3774	102	26644	0.69	0.4	1.3	0.32
BARD1	22	3667	27	26078	5.81	3.3	10.4	1.67×10^{-8}
BRCA1	234	4731	82	26911	16.62	12.8	21.5	$<2.2 \times 10^{-16}$
BRCA2	97	4731	109	26791	5.08	3.8	6.7	$<2.2 \times 10^{-16}$
BRIP1	15	3667	49	26840	2.24	1.2	4.0	0.01
CDH1	1	4733	3	25961	1.83	0.1	16.6	0.49
CDKN2A	1	1040	9	24312	2.60	0.1	18.9	0.34
CHEK2	13	3764	232	25215	0.37	0.2	0.7	1.19×10^{-4}
MLH1	-	1975	10	26639	ND	ND	ND	ND
MRE11A	2	3667	25	26767	0.58	0.1	2.3	0.77
MSH2	-	1975	6	25329	ND	ND	ND	ND
MSH6	5	1975	34	26151	1.95	0.7	4.8	0.19
NBN	9	3667	41	26264	1.57	0.7	3.2	0.20
NF1	9	3436	29	26130	2.36	1.1	5.0	0.04
PALB2	66	3933	30	26869	15.15	9.8	23.6	$<2.2 \times 10^{-16}$
PMS2	3	1975	32	26230	1.25	0.3	3.9	0.73
PTEN	1	4848	1	24166	4.99	0.1	191.9	0.31
RAD50	7	3667	58	26474	0.87	0.4	1.9	0.85
RAD51C	13	3667	37	26647	2.56	1.3	4.8	7.04×10^{-3}
RAD51D	8	3436	8	26555	7.74	2.9	20.7	1.62×10^{-4}
TP53	8	4865	18	26789	2.45	1.0	5.6	4.95×10^{-2}
TP53†	6	1008	18	26789	8.88	3.5	22.5	1.73×10^{-4}

(-) No mutated alleles; ND = not determined; OR =Odds ratio; CI = confidence interval

* Significance of associations were estimated using Fisher's exact test. All tests were two-sided.

†Age at diagnosis of 40 years or younger

Supplementary Table 8: Associations with TNBC by gene for patients of all race/ethnicity in the clinical cohort and TNBCC studies

Gene	Clinical Cohort		TNBCC Cohort		ExAC controls		Clinical TNBC Risk				TNBCC TNBC Risk			
	Mutated Alleles	No. Cases	Mutated Alleles	No. Cases	Mutated Alleles	No. Control	OR	95%CI lower	95%CI upper	p-value*	OR	95%CI lower	95%CI upper	p-value*
ATM	16	6556	4	2143	185	52161	0.69	0.41	1.15	0.18	0.53	0.18	1.40	0.26
BARD1	44	6370	10	2143	48	50931	7.35	4.80	11.17	$<2.2x10^{-16}$	4.96	2.41	9.83	$1.01 x10^{-4}$
BRCA1	431	8411	163	2143	121	52629	22.83	18.60	28.03	$<2.2x10^{-16}$	34.37	27.06	43.67	$<2.2x10^{-16}$
BRCA2	195	8411	56	2143	203	52344	6.04	4.95	7.36	$<2.2x10^{-16}$	6.81	5.03	9.17	$<2.2x10^{-16}$
BRIP1	26	6370	9	2143	89	52528	2.41	1.52	3.76	$2.35 x10^{-4}$	2.48	1.21	4.93	0.01
CDH1	5	8383	ND	ND	11	50853	2.76	0.92	7.85	0.06	ND	ND	ND	ND
CDKN2A†	4	1763	ND	ND	14	47527	7.71	2.34	24.10	$3.34 x10^{-3}$	ND	ND	ND	ND
CHEK2	20	6543	2	2143	401	49679	0.38	0.24	0.59	$1.31 x10^{-6}$	0.12	0.02	0.43	$1.33 x10^{-5}$
MLH1	4	3438	ND	ND	18	52066	3.37	1.05	9.64	0.04	ND	ND	ND	ND
MRE11A	6	6370	4	2143	39	52368	1.27	0.52	2.96	0.63	2.51	0.82	6.84	0.09
MSH2	2	3438	ND	ND	7	49369	4.10	0.61	18.16	0.11	ND	ND	ND	ND
MSH6	9	3438	1	372	116	51038	1.15	0.56	2.24	0.58	1.18	0.06	6.6	0.57
NBN	12	6370	2	2143	66	51539	1.47	0.78	2.77	0.21	0.73	0.13	2.70	1.00
NF1	9	6004	ND	ND	49	51239	1.57	0.75	3.23	0.20	ND	ND	ND	ND
PALB2	102	6877	20	2143	71	52529	11.05	8.14	15.09	$<2.2x10^{-16}$	6.93	4.12	11.50	$3.15 x10^{-10}$
PMS2	3	3438	ND	ND	62	51336	0.72	0.19	2.20	0.80	ND	ND	ND	ND
PTEN	3	8593	1	2143	5	48203	3.37	0.70	13.79	0.11	4.50	0.19	32.21	0.23
RAD50	11	6370	6	2143	141	51949	0.64	0.32	1.16	0.19	1.03	0.44	2.32	0.83
RAD51C	25	6370	8	2143	69	52206	2.97	1.86	4.75	$1.71 x10^{-5}$	2.83	1.25	5.86	0.01
RAD51D	13	6002	7	2143	26	51896	4.33	2.19	8.66	$9.50 x10^{-5}$	6.53	2.74	14.74	$2.64 x10^{-4}$
TP53	12	8615	2	2143	29	52214	2.51	1.21	4.98	0.01	1.68	0.29	6.35	0.35
TP53 ‡	6	1864	2	526	29	52214	5.80	2.35	13.65	$1.15 x10^{-3}$	6.86	1.16	25.98	0.04
XRCC2	ND	ND	2	2143	30	52539	ND	ND	ND	ND	1.64	0.28	6.14	0.36

CI= confidence interval; OR = odds ratio; ND = not determined;

* Significance of associations were estimated using Fisher's exact test. All tests were two-sided.

† Excluding common variant p.Ile49Thr

‡ Age at diagnosis of 40 years or younger

Supplementary Table 9: Associations with TNBC excluding TNBC with personal and family history of ovarian cancer

Gene	Clinical Cohort		ExAC controls		TNBC Risk			
	Mutated Alleles	No. Cases	Mutated Alleles	No. Controls	OR	95%CI lower	95%CI upper	p-value†
ATM	9	3414	102	26644	0.69	0.34	1.35	0.37
BARD1	21	3313	27	26078	6.14	3.39	11.09	1.27×10^{-8}
BRCA1	176	4270	82	26911	13.79	10.56	18.01	$<2.2 \times 10^{-16}$
BRCA2	86	4270	109	26791	4.99	3.73	6.67	$<2.2 \times 10^{-16}$
BRIP1	12	3313	49	26840	1.99	1.03	3.79	0.04
CDH1	1	4277	3	25961	2.02	0.08	18.35	0.46
CDKN2A	1	988	9	24312	2.74	0.13	19.89	0.33
CHEK2	13	3402	232	25215	0.41	0.23	0.73	7.02×10^{-4}
MLH1	1	1720	10	26639	1.55	0.07	10.53	0.50
MRE11A	1	3313	25	26767	0.32	0.02	1.87	0.35
MSH2	1	1720	6	25329	2.46	0.11	18.31	0.37
MSH6	2	1720	34	26151	0.89	0.15	3.57	1.00
NBN	7	3313	41	26264	1.35	0.60	3.07	0.49
NF1	7	3093	29	26130	2.04	0.87	4.75	0.10
PALB2	59	3557	30	26869	14.97	9.53	23.32	$<2.2 \times 10^{-16}$
PMS2	1	1720	32	26230	0.48	0.02	2.93	0.72
PTEN	2	4379	1	24166	11.04	0.86	320.08	0.06
RAD50	3	3313	58	26474	0.41	0.11	1.27	0.15
RAD51C	9	3313	37	26647	1.96	0.93	4.07	0.09
RAD51D	6	3092	8	26555	6.45	2.23	18.52	1.83×10^{-3}
TP53	6	4395	18	26789	2.03	0.79	5.15	0.14
TP53†	4	854	18	26789	6.98	2.17	19.96	4.25×10^{-3}

OR = Odds ratio; CI = confidence interval

* Statistical significance of associations were estimated using Fisher's exact test. All tests were two-sided.

† Age at diagnosis of 40 years or younger

Supplementary Table 10: Associations with TNBC excluding personal and family history of colorectal cancer

	Clinical cohort		ExAC controls		TNBC Risk			
Gene	**Mutated Alleles**	**No. Cases**	**Mutated Alleles**	**No. Controls**	**OR**	**95%CI lower**	**95%CI upper**	**p-value***
ATM	9	2901	102	26644	0.81	0.40	1.59	0.63
BARD1	17	2807	27	26078	5.86	3.06	10.76	3.13×10^{-7}
BRCA1	188	3637	82	26911	17.39	13.39	22.61	$<2.2 \times 10^{-16}$
BRCA2	88	3637	109	26791	6.01	4.51	8.01	$<2.2 \times 10^{-16}$
BRIP1	12	2807	49	26840	2.34	1.21	4.48	0.01
CDH1	-	3675	3	25961	ND	ND	ND	ND
CDKN2A	2	665	9	24312	8.14	1.26	36.69	0.03
CHEK2	13	2895	232	25215	0.49	0.27	0.85	8.07×10^{-3}
MLH1	-	1306	10	26639	ND	ND	ND	ND
MRE11A	2	2807	25	26767	0.76	0.13	2.96	1.00
MSH2	-	1306	6	25329	ND	ND	ND	ND
MSH6	3	1306	34	26151	1.77	0.46	5.45	0.26
NBN	5	2807	41	26264	1.14	0.43	2.90	0.80
NF1	6	2626	29	26130	2.06	0.84	4.85	0.13
PALB2	55	3030	30	26869	16.40	10.50	25.77	$<2.2 \times 10^{-16}$
PMS2	1	1306	32	26230	0.63	0.03	3.86	1.00
PTEN	2	3737	1	24166	12.94	1.01	375.09	0.05
RAD50	3	2807	58	26474	0.49	0.13	1.50	0.28
RAD51C	11	2807	37	26647	2.83	1.41	5.68	4.73×10^{-3}
RAD51D	6	2624	8	26555	7.60	2.63	21.83	8.38×10^{-4}
TP53	5	3744	18	26789	1.99	0.71	5.48	0.19
TP53†	4	767	18	26789	7.78	2.41	22.23	2.93×10^{-3}

(-) No mutated alleles; ND = not determined; OR = Odds ratio; CI = confidence interval

* Significance of associations were estimated using Fisher's exact test. All tests were two-sided.

† Age at diagnosis of 40 years or younger

Supplementary Table 11. Estimated risks of TNBC associated with pathogenic variants in TNBC predisposition genes among African American TNBCs

Gene	Clinical cohort		gnomAD AFR controls		Case-control comparison			
	Mutated Alleles	Cases	Mutated Alleles	Controls	OR	95%CI lower	95%CI upper	p-value*
BARD1	10	873	7	7450	12.25	4.60	35.23	1.53×10^{-6}
BRCA1	65	1237	9	7641	45.82	23.06	94.76	3.35×10^{-16}
BRCA2	38	1237	20	7590	11.82	6.88	20.77	2.97×10^{-16}
BRIP1	4	873	19	7637	1.84	0.57	5.59	0.29
PALB2	7	948	15	7639	3.77	1.44	9.77	7.54×10^{-3}
RAD51C	4	873	2	7622	17.49	3.26	130.50	1.41×10^{-3}
RAD51D	2	846	2	7596	8.99	0.97	83.12	5.25×10^{-2}
TP53	1	1250	1	7633	6.11	0.16	235.14	0.26

OR = Odds ratio; CI = confidence interval; AFR = ExAC African controls

* Significance of associations were estimated using Fisher's exact test. All tests were two-sided.

Supplementary Table 12. Frequency of mutations by age at TNBC diagnosis and family history of breast and ovarian cancer among African American TNBC patients from the clinical cohort

Family cancer history	Age at TNBC diagnosis									
	<35 years		35-39 years		40-49 years		50-60 years		>60 years	
	PV carriers	%PV	PV carriers	%PV	PV carriers	%PV	PV carriers	%PV	PV carriers	%PV
No breast, no ovarian										
BRCA1	6	12.2	3	5.7	6	3.7	2	1.5	0	0.0
Other TNBC genes	1	2.0	3	6.0	8	5.9	6	5.2	2	7.2
All TNBC genes total	7	14.3	6	11.7	14	9.6	8	6.6	2	7.2
Other breast cancer	0	0.0	0	0.0	2	1.2	0	0.0	0	0.0
All breast cancer genes	7	14.3	6	11.7	16	10.8	8	6.6	2	7.2
One relative with breast, no ovarian										
BRCA1	11	26.2	3	7.5	5	4.6	2	1.4	0	0.0
Other TNBC genes*	2	6.2	1	2.5	10	10.7	9	7.5	3	6.6
All TNBC genes total	13	32.4	4	10.0	15	15.3	11	8.9	3	6.6
Other breast cancer†	1	2.3	0	0.0	0	0.0	0	0.0	1	2.0
All breast cancer genes	14	34.7	4	10.0	15	15.3	11	8.9	4	8.5
≥2 relative with breast, no ovarian										
BRCA1	11	57.9	1	6.7	7	13.2	3	3.8	3	6.5
Other TNBC genes*	2	10.5	1	6.7	2	4.3	5	7.3	1	2.2
All TNBC genes total	13	68.4	2	13.3	9	17.5	8	11.1	4	8.7
Other breast cancer†	0	0.0	0	0.0	0	0.0	0	0.0	1	2.3
All breast cancer genes	13	68.4	2	13.3	9	17.5	8	11.1	5	11.0
Any relative with ovarian										
BRCA1	2	22.2	1	33.3	2	5.9	0	0.0	1	5.0
Other TNBC genes*	1	11.1	0	0.0	4	13.2	3	7.8	1	7.7
All TNBC genes total	3	33.3	1	33.3	6	19.1	3	7.8	2	12.7
Other breast cancer†	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
All breast cancer genes	3	33.3	1	33.3	6	19.1	3	7.8	2	12.7

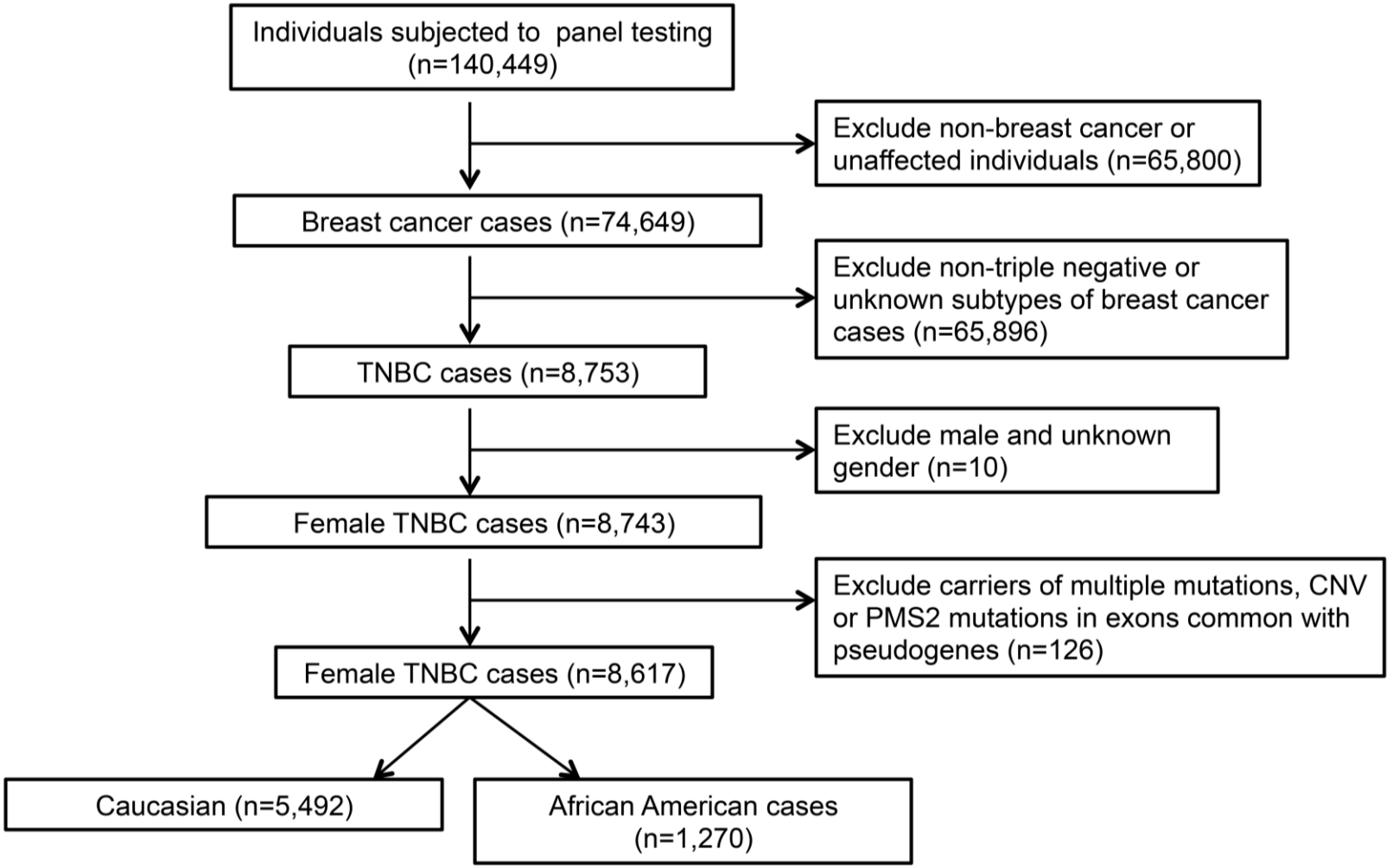
PV = pathogenic variant; %PV = % cases with pathogenic variants

*Other TNBC genes: *BARD1, BRCA2, BRIP1, PALB2, RAD51C, RAD51D, TP53*

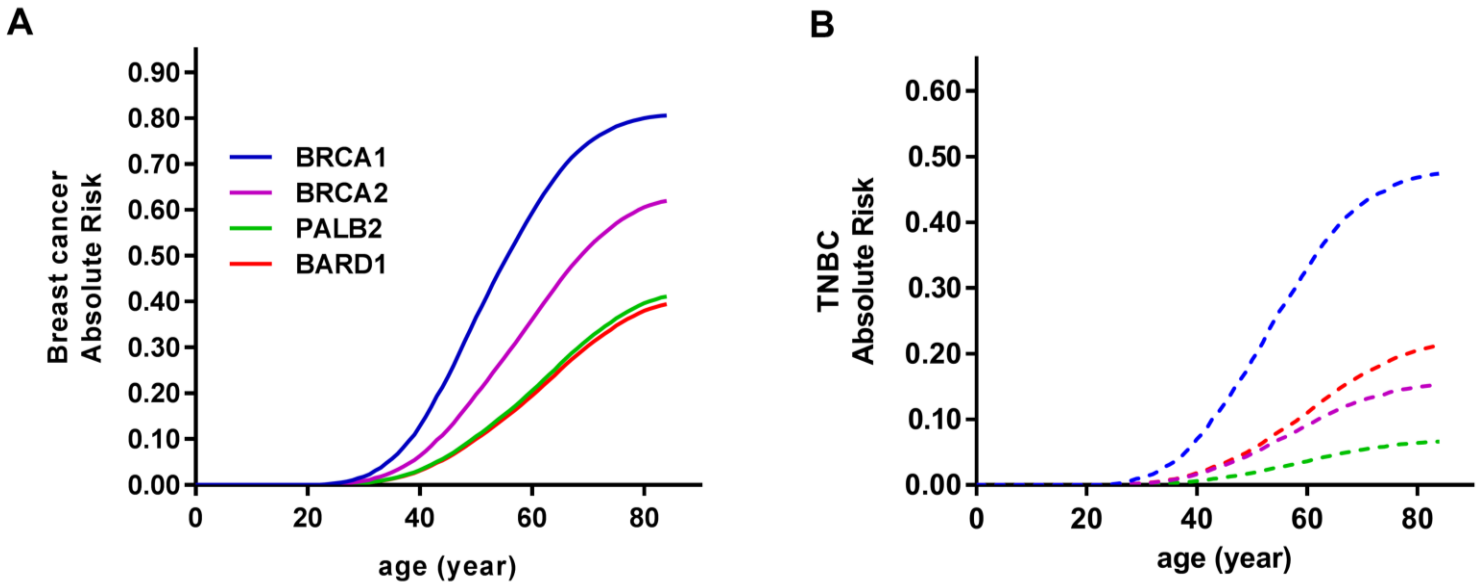
†Other breast cancer genes: *ATM, CHEK2, CDH1, PTEN*

Supplementary Figures

Supplementary Figure 1. Data cleaning and filtering for the TNBC clinical cohort



Supplementary Figure 2. Absolute risk estimates to age 85 for African-American overall breast cancer and triple negative breast cancer (TNBC)



A. Age-related risk curves for overall breast cancer for four genes are shown as color lines. B. Age-related risk curves for TNBC for four genes are shown as colored dashes.